

Chi Young Oh. Efficient Product Search Tools for Online Consumers with Little Product Knowledge: Clustering Versus Faceted Search. A Master's Paper for the M.S. in I.S degree. April, 2010. 58 pages. Advisor: Dr. Gary Marchionini

This study compares and contrasts different types of search tools—clustering search, rich-faceted search, and less-faceted search—in terms of their performance in online shopping product searches. In the experiment, two groups of people, a music-expert group and a computer-expert group, conducted searches for two different products (one for a novice search and the other for an expert search). Search performance was measured on three different shopping search websites—Clusty Shopping, PriceGrabber, and Yahoo Shopping. In addition, the participants' perceptions of each website's usefulness, directability/controllability, and their likelihood of future use are measured through their subjective ratings in questionnaires. The results of this research provide insight into how online consumers with different levels of product knowledge experience different types of search tools, with regard to search performance, perceived usefulness and controllability, and website preference.

#### Headings:

Use studies / Electronic commerce

Internet search engines / Evaluation

Classification / Systems / Faceted

Classification / Automation

EFFICIENT PRODUCT SEARCH TOOLS FOR ONLINE CONSUMERS WITH  
LITTLE PRODUCT KNOWLEDGE: CLUSTERING VERSUS FACETED SEARCH

by  
Chi Young Oh

A Master's paper submitted to the faculty  
of the School of Information and Library Science  
of the University of North Carolina at Chapel Hill  
in partial fulfillment of the requirements  
for the degree of Master of Science in  
Information Science.

Chapel Hill, North Carolina

April 2010

Approved by

---

Dr. Gary Marchionini

## Table of Contents

Introduction.....	2
Literature Review.....	10
Online Consumer Information Search and its Tools.....	11
Subject Knowledge .....	20
Method .....	21
Research Method and its Appropriateness for This Research .....	21
Sampling .....	22
Data Collection Instruments and Materials.....	24
Study Procedures .....	25
Result .....	27
Data Analysis .....	27
Search Performance .....	28
Perceived Values: Usefulness and Directability/Controllability .....	32
Usefulness .....	32
Directability/Controllability.....	33
Within Websites.....	36
Likelihood of Future Use.....	37
Further Comparison of Performance and Perceived Values.....	38
Performance and Likelihood of Future Use.....	38
Perceived Usefulness and Likelihood of Future Use.....	38
Perceived Directability and Likelihood of Future Use .....	39
Discussion .....	40
Advantages and Limitations of the Study .....	47
Conclusions.....	48
References.....	51
Appendices.....	53
Acknowledgements.....	56

## **Introduction**

Use of the Internet has changed many of our patterns and ways of life. Online shopping is one change that significantly affects our way of life today. Instead of driving a car to go to a store, walk around and shop, now consumers can go to a variety of virtual stores from within their homes and access almost all of the products that are sold in local stores. If they put forth some effort and navigate through several online stores, they may find better deals on products compared to what they could find in their local stores. The same navigation performed offline may lead to fatigue and more time spent. Those efforts could cost so much to consumers that they might not want to put the same amount of effort offline that they may otherwise have considered if online. Shopping can be done relatively easily online with lower search costs with the assistance of web search engines that provide users with a variety of product search results for their online shopping at any time.

Oftentimes, however, numerous product results on shopping search engines can also cause inconvenience in searching for the right product when consumers have to find what they want among the hundreds or thousands of product search results. Smith, Bailey, & Brynjofsson (1999) pointed out that search costs might decrease in online markets, but consumers may face information overload problems with many different search results. Online consumers may need other helpful tools or aids for their product search. Searching for the right product may be somewhat easier if consumers have significant knowledge

about the product because then they may be able to form more appropriate queries or select the right categories from the options that the product search websites provide. But for individuals who have little product knowledge, knowing where to find the products online could be relatively more difficult. To online consumers, especially those who do not have much product knowledge about the product, reliable and systematic assistance that supports their navigation in the sea of the search results may be helpful. Today there are many shopping search websites available online which provide some useful search assistance features that were designed to help users refine or organize their search results. According to the findings by Jaillet (2003), novice consumers with little product knowledge are more likely to rely on readily available data on the search engines and portals, whereas consumers who are experts in the product are more likely to go directly to more specialized websites. The findings may imply that consumers with little product knowledge are more likely to get help from shopping search engines than expert consumers with more product knowledge, and the helpful search features in the shopping search engines may better help novice consumers than expert consumers.

Most search engines employ certain search technology features that help their users perform better information searches. Among them, faceted searches and clustering searches are major search technology features that are used in search systems to help users locate information better, especially to help them to group and refine the search results more effectively. A facet is a dimension or feature type related to information, and a faceted search interface builds a set of category hierarchies consisting of different facets (or feature types) obtained from the information of the search results. It helps users refine and explore the search results by allowing users to click on the facets made from

available information of the search results. For example, in a shopping search, the facets can be price range, brands, performance levels, or other major specifications. The faceted search system is used in major shopping search systems such as Yahoo Shopping and Google Product Search.

Another newly studied search technology that helps users perform better searching and refining is clustering search. Clustering is a search technology which provides grouping among the search results to organize the results by meaningful topical clusters. Clustering features are used in some search engines such as Clusty.com and its shopping search, WebClust.com, and Carrot2.org. Kumar and Lang (2007) showed that clustering support provided by search engines helped to increase relevance and usefulness of search results especially when underspecified search terms were used for the search. When it comes to online shopping searches, under-specified terms were more likely to be produced by consumers who have less knowledge about the product, while well-specified terms were more likely to be formed by consumers with expert product knowledge. Thus it can be derived that the clustering feature of search engines can assist searching performance of consumers who have little product knowledge. Also, this implies that search assistance features providing users with some organizing or result refining support may help consumers with little product knowledge to find the specific products that they are looking for.

However, it is questionable which type of search tool can better help online consumers with little product knowledge locate relevant products. The clustering feature can help consumers' information seeking by providing automatic, on-the-fly, topical grouping of the product search results, whereas faceted searches can help users to refine

the product search results as they select pre-organized facets of information related to the products. This study will examine the effects of the two different search assistance tools—faceted searches and clustering searches—used in shopping search websites in terms of their helpfulness for online consumers with little product knowledge to find specific types of products more effectively. The effects of different search tools on online shoppers' search efficiency and the effects of the search tools on consumers' perception of usefulness and directability/sense of control when experiencing shopping will be investigated. The following are the research questions to be investigated in this study:

- Which type of search tool—clustering or faceted search—better assists online consumers possessing little product knowledge to effectively find the right products?
- Which type of search tool—clustering or faceted search—provide online consumers possessing little product knowledge with a better shopping search experience?

To answer this study's research questions, experiments with participants from two different backgrounds were conducted. They were asked to search for two products — one that they have background/expertise in and the other that they do not. During the experiments, participants' search logs for certain tasks on certain shopping search websites were recorded. Additionally, the participants' perceptions of the usefulness and directability/controllability (how well they were able to direct or control their searching and refining as they wanted to) of their search experiences were measured with subjective ratings through questionnaires.

Three different shopping search websites with different types of search tools have been selected. Shopping search websites that are specialized in certain product categories such as computer products (e.g., [www.newegg.com](http://www.newegg.com)) or certain product brands' online stores were excluded. Product search websites employing faceted searches were differentiated into two levels—one with rich-faceted search support and the other with less-faceted search support, thus two faceted search systems and one clustering supported system were sought for the experiments. When defining rich or less faceted support, numbers of relevant facets for the search tasks of this research were considered as criteria. Numbers of relevant facets that involves products' internal values, such as product types, functions, prices, and specifications were compared. Numbers of other facets including products' manufacturer, brand, or stores were excluded for the criteria since they may not affect the searching/refining performance of the tasks that the present study uses. The search tasks include searching for two items—a portable musical keyboard and an external hard drive. For both search tasks, Yahoo Shopping was considered as having fewer facets (one relevant facet for the musical keyboard task and four relevant facets for the hard drive task), while PriceGrabber was considered as having rich facets (two relevant facets for the musical keyboard task and eight relevant facets for the hard drive task) at the time of the research experiments. Thus, firstly, Yahoo Shopping (<http://shopping.yahoo.com/>) was selected as the shopping search website with less-faceted search support for the present study tasks. Secondly, PriceGrabber (<http://www.pricegrabber.com/>) was selected as the shopping search website with rich-faceted search support for the present study tasks. Thirdly, Clusty Shopping (<http://shopping.clusty.com/>), which is a clustering search system that provides general



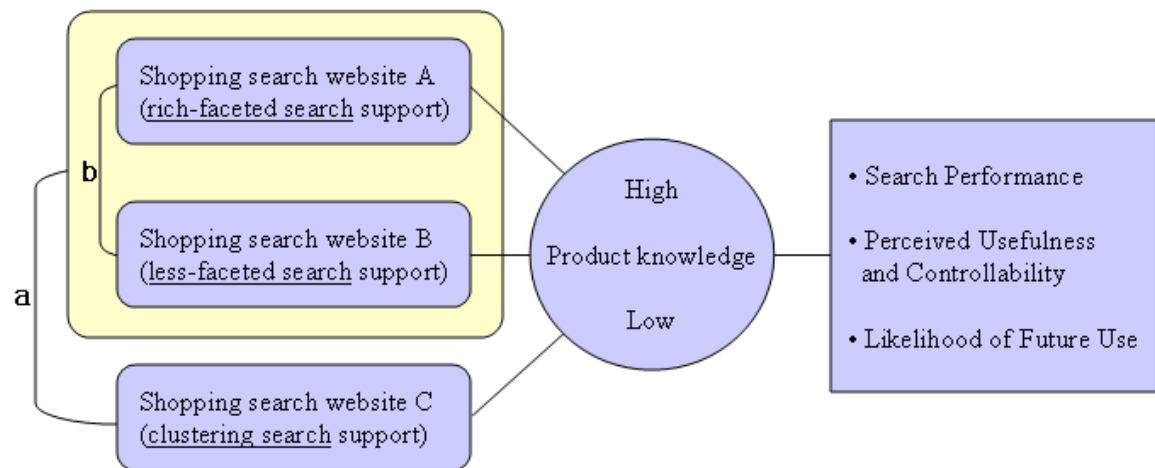
shopping search services, was selected as the shopping search website with clustering support.

Efficiency of clustering search was proven or acknowledged in some research (Kumar & Lang, 2007; Hearst, 2006b). Clustering can be useful when users need to clarify and narrow down a vague query, because it shows users the major themes of the search results (as cited in Hearst, 2006b). But its actual usefulness in shopping product searches is questionable, since it does not provide well-structured search aids that are predictable and have intuitive hierarchies for users to refine the search results accordingly. It may result in effective search performance, but users may find the various clusters confusing or hard to direct or control their searches with the help of clusters. However, in faceted searches, users can select facets that were pre-organized to refine the product search results and have a relatively greater sense of control over their searches, and their search performance may be higher than in clustering searches, since users can drill down the search results by using faceted searches that are predictable and intuitively structured.

Consumers who have little product knowledge may experience more search difficulties and need to expend more effort in navigating the search results across the search websites. Considering the understandable hierarchies of search results and the capability to help users to build up a complex query through navigation, which was addressed in Hearst (2006b), the faceted search was hypothesized to better help consumers with little product knowledge to find the right products effectively and to provide them with search experiences of better usefulness and sense of control (directability/controllability) than the clustering search. Also, the richness of the facets

used in the faceted search was considered to affect the search performance and subjects' perception of its usefulness and directability/controllability. Thus, the research model and the hypotheses for this study were formulated as follows:

#### ▪ Research Model



“a” relationship: comparison between faceted systems and clustering system

“b” relationship: comparison between rich-faceted system and less-faceted system

#### ▪ Variables

Independent variables: Different types of shopping search tools

- System A. Rich-faceted search support (PriceGrabber)
- System B. Less-faceted search support (Yahoo Shopping)
- System C. Clustering search support (Clusty Shopping)

Moderator variable: Consumers' product knowledge of the searching item

- High (Expert product knowledge)
- Low (Novice product knowledge)

Dependent variable: Search performance and user experience

- Search performance (Duration time to find the specific type of products)

-Usefulness and directability/controllability (Values perceived by consumers for their search experiences)

-Likelihood of future use of the website

▪ Hypotheses

H1a. Consumers with little product knowledge will perform the search task more efficiently on the shopping search websites with faceted search support than on the shopping search websites with clustering support. (Websites with faceted search support: system A and B)

H1b. Consumers with little product knowledge will perform the search task more efficiently on the shopping search website with rich faceted search support than on the shopping search website with less faceted search support.

H2a. Consumers with little product knowledge will find the search experience at the faceted search websites more useful and controllable than at the search website with clustering support.

H2b. Consumers with little product knowledge will find the search experience at the rich-faceted search websites more useful and controllable than at the less-faceted search website.

H3. Search performances for finding novice products will have a greater impact on consumers' willingness to use the website in the future than search performances for finding expert products.

H4. Consumers' perceived values (usefulness and directability/controllability) regarding the websites in novice product searches will have a greater impact on consumers' willingness to use the websites in the future than their perceived values of the websites in expert product searches.

The rest of this paper is organized as follows: a review of the literature related to the present topic, a description of the research methods for comparing the performance and effectiveness of the search tools in shopping search setting, analysis of the results, discussion and conclusion.

### **Literature Review**

This research is specifically concerned with the performance and users' experience of shopping search tools depending on users' subject knowledge related to the product. DeSarbo and Choi (1999) observed that the scope of consumer information search behavior literature is extremely vast and diverse, ranging from marketing and economics to psychology and sociology. Addressing literature in those disciplines would be overly ambitious and far beyond the scope of the present study. Rather, the scope of this study is within the search performance of search tools used on online shopping search websites, their effectiveness for users in terms of users' level of product knowledge, and the effects of using different search tools on the consumers' online

shopping experience. Two major streams of literature from the business discipline and the ILS (Information and Library Science) discipline are particularly relevant to the present topic: 1) online consumer information search and its tools, and 2) effects of subject knowledge on users' information search. Online shoppers' information search behavior is a part of the research interests in consumer behavior or marketing research, but search performance of the shopping search engines is not a major concern of consumer behavior or marketing research. Most of the studies concerning search performance of shopping websites were from information retrieval in the ILS research field and information systems / e-commerce / decision making research disciplines.

Relevant literature across those disciplines will be briefly reviewed for the two major streams.

### **Online Consumer Information Search and its Tools**

Jaillet (2003) investigated online shoppers' information search patterns and the search tools they use. When investigating users' patterns and website preferences for online shopping, Jaillet (2003) used consumers' product knowledge level as the independent variable. The author chose "bicycle" and "slide scanner" as the products to research and purchase. To find subjects who were experts in bicycle products, the author recruited 43 people from local cycling clubs. After going through the manipulation check test for measuring their product knowledge levels, 37 out of 43 subjects who were experts in bicycles but not in slide scanners were chosen for the experiment. The subjects were identified as having expert product knowledge about bicycles and novice product knowledge of slide scanners. The author explored the search behaviors and patterns of these 37 subjects as they searched online for information concerning the purchase of the

two products. The study results found a significant difference in users' information search patterns when they performed expert search tasks and when they performed novice search tasks. In the research outcomes, significant difference was also observed between product experts and product novices in the use of certain types of websites. The author analyzed these results and claimed that individuals with expert product knowledge make use of more specific websites specialized for the product category, such as niche portals or niche vendors' websites, while individuals with novice product knowledge rely on search engines, general portals, and general vendors' websites to find their desired product information. Considering these findings, search engines/general portals could be better designed in favor of online shoppers who have novice product knowledge rather than online shoppers who have expert product knowledge. The search websites would be able to provide more benefit to individual shoppers with novice product knowledge if their search services are designed to better support those users. This idea supports the basic assumption of the present study that individuals with novice product knowledge seek more help from search engines than individuals with expert product knowledge do. Jaillet (2003) classified the type of websites into comparison sites, general portals, search engines, general vendor's websites, and niche portal. However, the research did not classify the websites by different types of search tools. It did not consider the effects of the features used in those websites on online consumers' preferences for the websites. Online shoppers' actual product seeking on the shopping websites was not specifically examined through the research, but general information searching behavior for product purchase was observed.

Haubl and Trifts (2000) studied the effect of interactive consumer aid tools in online shopping environments. One tool was called Recommendation Agent, which assists consumers in the initial screening of the alternative products by recommending a set of products that are likely to be attractive to the consumer based on shopper's preference information. The other tool is Comparison Matrix, which assists consumers in doing in-depth comparison among alternative products. For the experiment, eighty undergraduate psychology students were recruited, and a controlled experiment was conducted to test the hypotheses about the effect of Recommendation Agent and the Comparison Matrix on the amount of search for product information, forming of consideration set, and decision quality. The task in the experiment was shopping for a product in each of two product categories –backpacking tents and compact stereo systems in an online store. The results showed that the Recommendation Agent provided significant assistance in consumers' product searches. The assistance included screening out a significant amount of irrelevant product information and facilitating in-depth comparisons among products, which allowed consumers to put less effort into searching products and making decisions, and increased confidence in their choices. The impact of the other search aid tool, Comparison Matrix, on consumers' product consideration was found to be very substantial, and it decreased the average number of alternative products which were seriously considered for purchase. The research focused on the effects of interactive decision aids used within a particular merchant's online store. Due to the similarities in the structure and function between online stores and shopping search engines, the research findings could imply that online shoppers would also receive significant support from search aid features in general shopping search engines.

Considering the effects of those search aid features addressed in their study, it could also be assumed that there are some other possible search-aid tools that can provide assistance to online shoppers to improve their shopping search performance. The focus of Haubl et al. (2000) was not about performance of the interactive aid tools used in product search or their effectiveness, but about the effect of the aid tools on online shoppers' consideration on alternative products to purchase and their subsequent decision making process.

Kumar, Lang, and Peng (2005) studied consumer search behavior in online shopping environments by investigating behavioral factors and technological factors which affect online shoppers' search performance. Specifically, the study examined the effects of independent variables including user ability, search engine capability, complexity of search tasks, and search heuristics on the dependent variable—online shopping search performance—which consisted of search cost and user satisfaction. The independent variable “user ability” was defined as users' experiences in using search engines and their subject knowledge on the object of the search. The hypothesis 1 was that stronger user ability leads to lower search cost, defined as being more effective and quicker in formulating search queries, processing search results and finding needed products. Search engine capability was defined as the functions or features of the search engines to improve end users' usability and effectiveness in searching. Many factors that affect search engine capability were discussed, including search interfaces that enhance usability, use of personalized features based on account log-in, visual display of product information, and search results refining function using topical clustering. Among them, use of the clustering feature was chosen as the independent variable, and it was hypothesized that higher search engine capability leads to better search performance



(Hypothesis 2). Search tasks were divided in terms of the complexity of the search task: Simple tasks were fact-based, known-item searches and complex tasks were incompletely specified, open-answer searches. It was hypothesized that higher complexity of search tasks leads to higher search cost (Hypothesis 3). Lastly, the study chose the level of restrictiveness in search heuristics as a moderating variable that affects the intensity of the relationships among the main variables. For example, the search heuristic, “find the first available deal,” was of weak restrictiveness, while the search heuristic, “find the best deal available,” was of strong restrictiveness. It was hypothesized that the impacts of user ability, search engine capability, and search task on search costs will be stronger when the employed search heuristic is more restrictive. During the experiment, the subjects’ queries, click streams, screens, web pages, and time data were collected. To measure the search cost from the time data, both the time that the subjects spent on each visited page and the total amount of time they spent to complete the assigned search tasks were recorded. By doing the web log analysis for subjects’ online behavior, the dependent variable, number of queries (search performance), was measured according to each independent variable: search task complexity, search engine capability, and search heuristic. The result showed that the differences in the independent variables led to the differences in the dependent variables, in the directions proposed in the hypothesis 2 and 3, although these results would not be regarded statistically significant due to the small sample size. In addition to the small sample size, there was another shortcoming regarding the study design. User ability, especially users’ domain knowledge of the searching task, was employed as another independent variable, but the authors did not differentiate subjects’ level of subject knowledge into high and low level, but just

controlled subjects' level of user ability to be high. The research did not clearly investigate whether online consumers with more product knowledge would show higher performance in their product searching and online consumers with little product knowledge would show lower performance in product searching. In addition, clustering was used as an independent variable for search engine capability, but the clustering feature is only one of the tools employed by search engines.. It would be desirable if more search-aid tools and features were used to test their impact on search performances.

Bruza, Dennis, and McArthur (2000) focused on comparing keyword-based search (Google), directory-based search (Yahoo) and phrase-based, query reformulation-assisted search (HiB – Hyper Index Browser) in terms of their search performance and effectiveness. The authors used eighteen queries generated to be a broad brushstroke of interesting Internet queries and measured the domain knowledge of the subjects about the queries. The research findings included that the directory-based search using Yahoo did not seem to provide increased relevance over the keyword-based search and the query-reformulating search, and that the query reformulating search (HiB) could significantly improve the relevance of the search results. However, interestingly, the results clearly indicated that the directory search engine (Yahoo), which provided a categorized information structure, performed better than the keyword search engine (Google) and the query reformulating search engine (HiB) when it came to shopping-related queries. In their research, users' search performance for shopping-related queries was not extensively investigated. The specific outcome of users' search performances on shopping-related queries could be further investigated in another research effort to find

whether additional search supports, such as directory search, faceted search, or clustering search, would help to improve users' shopping search performance.

Tools that assist online consumers' product searches include some conventional tools such as keyword search and directory search, and other newer tools such as product/keyword suggestion, clustering, and faceted hierarchical category. Kumar and Lang (2007) concluded in their research that search engines with clustering functions provide better support for online shoppers' searching performance compared to general keyword search engines without clustering support, when a search term was under-specified. Since it is more likely that consumers with little product knowledge would form under-specified terms for their queries, clustering feature of search engines may assist searching performance of consumers who have little product knowledge. But the study did not deal with other search tools such as faceted search that is being used in major online shopping search websites such as Yahoo Shopping and Google Product Search. It could be worthwhile to test the impact of other tools, such as faceted search, on online consumers' search performances and these tools' interplay with consumers' product knowledge.

According to Hearst (2006a), which presents design recommendations for hierarchical faceted search interface, hierarchical faceted search has been found to be a highly understandable model for search interfaces. Especially, many shopping websites and shopping search engines have employed faceted searches for their product search interface. In her usability experiments conducted for the Flamenco project, it was found that users felt comfortable when they used faceted searches and they felt they were in control while navigating multiple facet hierarchies simultaneously. The research claimed

that a key component to successful faceted search interfaces is the seamless integration of keyword search. The study suggested that in most cases users preferred the search results where their queries matched broader categories of information. The most probable reason for this preference was that, in that setting, users can organize the search results by using the faceted search interface. The research revealed the usefulness of a search system employing a faceted search interface. However, the research did not address the search performance of the search systems that use faceted search interfaces.

Faceted search interface was compared with a clustering search interface in terms of its usability in Hearst (2006b). The research described two popular information-grouping technologies—clustering and faceted categorization—and compared them with each other based on the results of usability studies. Clustering's characteristics of being fully automatable and easily applicable to any text collection were chosen as its greatest advantage. The article addressed that clustering helps users to clarify and sharpen a vague query, by guiding users to different themes among the search results. But clustering was found disadvantageous in that the grouping result is difficult to predict and its sub-hierarchy clusters are not intuitive. The author's usability test results showed that users do not like disorderly groupings such as the results of clustering, but prefer understandable hierarchies of search results. On the other hand, faceted category search was found to allow flexible ways to access the search results and help users to build up a complex query through navigation within the hierarchy of different facets and their sub-hierarchies. Hearst's usability studies revealed that, especially for browsing tasks, faceted search was highly preferred over the standard keyword-and-results listing search interfaces used with web search engines. But the shortcoming of the faceted category

search, as opposed to the clustering search, was that the categories of interest must be known in advance, and that so far, in most cases, the category hierarchies in faceted searches needed to be built manually. Automated category building for faceted searches has been only partly successful. The studies introduced in the article showed that a faceted category search was preferred to a clustering grouping search in terms of usability. However, the search performance of those two search systems was not addressed in the paper. It is still questionable whether a faceted search is superior to a clustering search in shopping product search performances. The research used just one type of faceted search system. However, the effects of the richness/number of facets in faceted search systems on the search performances could be further explored. Does use of more facets lead to higher search performance and higher satisfaction in users' search experience? What kinds of facets are more useful in certain search tasks or settings?

In the business literature, there has been little research about the effect of search engine capability on search performance except for some research on a clustering search and general keyword search tools. In the ILS field, there was a considerable amount of research about the search engines using a faceted search or a clustering search, but not on the relationship between those search interfaces and the search performance in terms of the search effectiveness or search efficiency in shopping search systems. Further, there has been little to no research conducted to investigate the effect of a faceted search or a clustering search in terms of their degree of capability (e.g. use of more facets – higher search engine capability) on search performance of users with different levels of product knowledge.

## **Subject Knowledge**

In previous research in the Information and Library Science field, domain knowledge was found to play a significant role in users' information search. Hsieh-Yee (1993) showed that subject expertise and search experience affected the use of search tactics of users. In the research, novice and experienced users searched questions in a database, and their use of search tactics were recorded by protocols, transaction logs, and observation, which were operationalized afterwards to describe the differences between the two groups of users. The study showed that subject expertise affected search performances particularly for experienced search engine users. The role of subject expertise in users' search performances needs to be further explored.

Other research investigated the effects of domain knowledge on users' searching performance. Bruza, Dennis, and McArthur (2000) hypothesized in their research that the search engines that provide some structure to the information domain might help users with insufficient domain knowledge, as they proceeded to compare the search performances of three different types of search engines. However, analysis of the results did not indicate significant differences between the performance of users with some domain knowledge and the performance of users with little domain knowledge. Their claim is against some of other research which revealed the relevant effects of domain knowledge on users' information search. Wildemuth (2004) found that if users had domain knowledge about the search context, it affected their search behaviors. In Wildemuth's study, the search tactics of medical students in searching a database in microbiology were examined, as the students answered problems and searched the database on three occasions (prior to taking the microbiology course, immediately after

the microbiology course, and six months after the end of the microbiology course) over a nine-month period. The research findings showed that low domain knowledge of users was associated with less efficient incorporation of concepts in the search and more errors in reformulating the search tactics. Although it investigated the effects of domain knowledge on users' search behavior in database searches, her research may indicate that users with low domain knowledge would have more difficulties in searching for certain information in search websites than users with high domain knowledge.

Bruza et al. (2000) did not indicate significant differences in search performances between users with some domain knowledge and users with little domain knowledge in general information searches. However, their findings contradict other research (Hsieh-Yee, 1993; Jaillet, 2003; Wildemuth, 2004) which found relevant effects of domain knowledge on users' information search. Subject knowledge could also be an important factor that affects users' information searches in a specific online environment such as online shopping search websites where many new search tools are increasingly used. Interdisciplinary, but specific research concerning the relationships between users' product knowledge and their search performance and experience when using different types of search tools, such as a faceted search or a clustering search, would be beneficial by contributing to the interdisciplinary research of e-commerce, marketing, and search interface design in this fast-growing online environment.

## **Method**

### **Research Method and its Appropriateness for This Research**

Transaction logs of subjects' search behavior will be recorded to gather data on subjects' product search and the search performance of each shopping search website.

Transaction logs are records of people's interaction with computer applications (Wildemuth, 2009). These records can be generated by using transaction log tracking software. Transaction logs can be data that contain the type, content, or time of users' information search. They can be time-stamped logs of subjects' mouse clicks or web-browser page changes. Transaction log analysis is commonly used in the information and library science field to collect users' online transaction and search data. By using log analysis, it is possible to know when users visited certain websites, what they explored on these websites, what keywords they used for searching, what hyperlinks they clicked on, and when they left the websites. The log analysis method is beneficial for use in this study because subjects' web search transaction records—including URLs, web page names, and especially time access stamps—can be obtained in a relatively easy, precise, and unobtrusive way (Jansen, Spink, & Taksa, 2009).

For the present study, subjects' transaction logs were used to record subjects' product search activities and included time stamps for each action as well as the beginning and end of each search task.

### **Sampling**

For this experiment, it is assumed that subjects are consumers who are going to buy two products—one for which they have relatively above-average product knowledge and the other for which they have little or no expertise. In the experiment design, the subjects were assumed to have certain preferences for the products, and thus general features and price ranges of the products were provided as search conditions. Slim, external hard drives and portable musical keyboards were chosen as the products for the experiment.



In order to conduct more powerful, extended comparisons of the results, a mixed study design which uses both within-subjects and between-subjects designs was chosen (Tullis, & Albert, 2008). The subjects consisted of two groups, one group of people who are familiar with computers and computer peripherals or are computer science majors, and the other group of people who have played the piano/musical keyboard before, have music backgrounds, or are music majors. To find those subjects, advertisement emails were sent out on listservs and printed advertisements were posted on bulletin boards in the computer science department, school of information and library science, and music department and other related organizations of local universities, specifically UNC-Chapel Hill and NC State University. In the advertisement, the experiment duration time was stated to be about 30 minutes, and \$6 gift card, as the compensation, was written on the advertisements to be provided to the participants. The advertisements specified that the participants should be familiar with online shopping websites and meet the subject qualifications—either for music group or for computer group, but “NOT BOTH.” Those who have relative expertise in computer and computer peripherals were considered as having more background knowledge or product knowledge about external hard disk drives, but relatively less product knowledge about musical keyboards, and that those who have relative expertise in music and piano were considered as having more background knowledge or product knowledge about musical keyboards, but relatively less product knowledge about external hard disk drives. 18 people were recruited from each group—the computer group and the music group—totaling 36 people. Before subjects participated in the experiment, they were asked questions as a manipulation check for their qualification. Participant’s product knowledge or recent experience about

both types of products and their eligibility was checked before the experiments began.

One person who was interested in participating in the research study was found to be irrelevant to either of the two groups and was not used for the study.

#### ▪ Subjects Groups

Computer Group. Those who have above-average knowledge and experiences about computer/computer peripherals, but have little background about music/piano.

Music Group. Those who have above-average knowledge and experiences about music/piano, but have little background about computer/computer peripherals.

### **Data Collection Instruments and Materials**

For search performance data collection, participants' web browser activities were recorded as log data, which provided precise measurements of subjects' search activities in an easy and unobtrusive manner. The data show when the subject began a product search task and when the task ended. Each subject's duration time for each search task on all three shopping search websites was obtained. Subjects' search behavior on the PC screen was recorded as well.

For the measurement of the quality of the product search experiences, questions for subjects' perceived usefulness of each website for each expert/novice search, how much they thought they could direct/control their searches, and how much likely they would like to use each search website for their future product searches were posed in a questionnaire. The participants were requested to fill out the questionnaire after they had

finished each of the expert/novice search tasks. Each answer choice was made on a 7-point Likert scale: 1 (very bad) – 2 (bad) – 3 (somewhat bad) – 4 (neutral) – 5 (somewhat good) – 6 (good) – 7 (very good) OR 1 (Not at all) – 2 – 3 – 4 (neutral) – 5 – 6 – 7 (very much so). The results were meant to reveal which product search system was perceived to be more useful when searching for certain products by certain groups of subjects in expert/novice search conditions.

### **Study Procedures**

First, detailed information and instructions for search tasks were prepared. The information included the order of tasks for each subject, context, explanations, and conditions about the products to search for, and the actions to do when subjects start and end each of the search tasks. Thirty-six participants, who were recruited through the advertisement and manipulation-checked before the experiment, attended the experiments. Total duration time of each experiment was about 30 minutes, varying across participants depending on how long they spent in searching for products and answering the questionnaire. Each of the subjects received a printed copy of instructions for the experiments, when they began the experiment (See appendix A). On the first page of the printed material, the order of the tasks at certain shopping search websites was provided. The order of shopping search websites was counterbalanced. From the three different shopping websites, six types of order ( $3! = 3 \times 2 \times 1$ ; A-B-C, A-C-B, B-A-C, B-C-A, C-A-B, C-B-A) were available. For products to search for, participants were first given the product that they have expertise in, and then the product that they do not have expertise in. Thus, the subjects were asked to search for products with their expertise on

the three shopping search websites first, and for products without their expertise on the three websites later. Each person's tasks involved two products and three shopping search websites.

In the printed instruction materials, products to search for were described with minimal information, such as product type and some features. Subjects were instructed to enter basic keywords for their initial searches—"hard drive" for external hard drive searches and "musical keyboard" for portable musical keyboard searches—and to refine the search results by using the search-aid features (facets or clusters) on the left side of the websites. They were asked to find 5 items which would satisfy specified product conditions and were allowed to use further modified keyword searches if they could not find enough items by refining the search results.

No time limit was set for the tasks to avoid any effect that a time restriction may have had on subjects' search behavior and heuristics (Kumar et al., 2005). However, the tasks were not supposed to take more than 30 minutes total (the entire search took about 20 minutes on average). Since participants knew the experiment would take about 30 minutes, they were able to have at least some feeling of time constraint, which is usually given to shoppers/consumers in many cases. After finishing their search tasks for each product, the subjects were to answer the questions in the following questionnaire. Two sets of the same questions were provided for their search tasks for each of the two products. The questions measure the participants' perception on the usefulness and directability/controllability for each of the two given product search tasks on the three

product search websites. It also measures the participants' subjective ratings for the likelihood of future use of each product search website.

## **Result**

### **Data Analysis**

Thirty-six people with either music expertise/background (18 people) or computer expertise/background (18 people) participated in the research experiments. The duration time for each of their search tasks was recorded and their answers to the questions about their search experiences during the tasks were obtained through the questionnaire. Each participant conducted search tasks for both musical keyboards and external hard drives. For participants in the music group, searches for musical keyboards were regarded as an "expert search" and searches for external hard drives were regarded as a "novice search," while searches for musical keyboards were regarded as a "novice search" and searches for external hard drives were regarded as an "expert search" for participants in computer group.

The duration time for each search task was used to calculate relevant statistics to test hypotheses 1 and 3 of this research. For the statistical testing, t-test and ANOVA were used to compare the Mean and variance differences between duration times obtained in the expert condition and the novice condition on the three different websites. Significance level of .05 was used. Also, the answers to the questions about subjects' perceived usefulness, directability, and likelihood of future use of the website were coded in an MS Excel sheet. The coded answers were used to calculate relevant statistic data to

compare the usefulness, directability, and likelihood of future use of the three search systems and to test the hypotheses 2 and 4. ANOVA and t-tests were used to test if there was a statistically significant difference in the subjects' perceptions on the usefulness, directability/controllability, and likelihood of future use among the three search systems. If a difference was found among the three search systems, statistical differences between each pair of systems were revealed. The statistical analysis results were used to verify whether the hypotheses would be supported. Lastly, the correlation coefficients (Spearman's rho) between search performances, perceived usefulness, directability, and likelihood of future use were calculated to test hypotheses 3 and 4 and to find out further implications. The results and implications will be discussed in the next section of the paper.

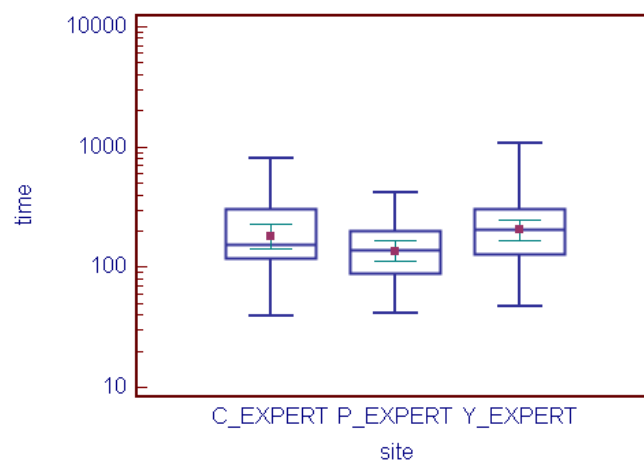
### **Search Performance**

Users' searches for their expert product were referred to as "expert" searches and their searches for novice product as "novice" searches. When participants did the "expert" searches (music group for musical keyboards and computer group for external hard drives), differences were found among the search duration times on the three websites. ( $P = 0.026$ ; Sig. level = 0.05). For comparisons of each pair, statistically significant differences were found in the pairs of PriceGrabber-Yahoo Shopping and PriceGrabber-Clusty Shopping, but not in the pair of Yahoo Shopping-Clusty Shopping. Significance values of the mean difference between each pair are shown in Table1.

Table1. Search duration times and mean differences for expert searches (unit: second;  
Mean)

<b>Clusty-Expert</b>	<b>PriceGrabber-Expert</b>	<b>Yahoo-Expert</b>
230.2	160.2	244.9

\*ANOVA test:  $P = 0.026$ , Sig. level = 0.05



<b>t-test</b>		<b>Clusty-Expert</b>	<b>PriceGrabber-Expert</b>	<b>Yahoo-Expert</b>
		230.2	160.2	244.9
<b>Clusty-Expert</b>	230.2		$P = 0.0357$	$P = 0.7223$
<b>PriceGrabber-Expert</b>	160.2	$P = 0.0357$		$P = 0.0143$
<b>Yahoo-Expert</b>	244.9	$P = 0.7223$	$P = 0.0143$	

This result shows us that in the expert search condition it took less time on PriceGrabber than on both Yahoo Shopping and Clusty Shopping, but there were no significant differences between the search duration times on Yahoo Shopping and Clusty Shopping.

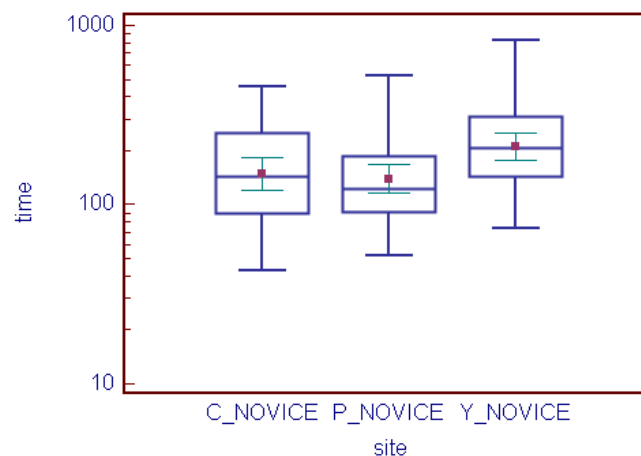
When participants did the “novice” searches (music group for external hard drives and computer group for musical keyboards), differences were also found among the search duration times on the three websites. ( $P = 0.016$ ; Sig. level = 0.05).

For comparisons of each pair, statistically significant differences were found in the pairs of PriceGrabber-Yahoo Shopping and Clusty Shopping-Yahoo Shopping, but not in the pair of PriceGrabber-Clusty Shopping. Significance values of the mean difference between each pair are shown in Table2.

Table2. Search duration times and mean differences for novice searches (unit: second; Mean)

<b>Clusty-Novice</b>	<b>PriceGrabber-Novice</b>	<b>Yahoo-Novice</b>
176.4	163.5	241.2

\*ANOVA test:  $P = 0.016$ , Sig. level = 0.05





<b>t-test</b>		<b>Clusty-Novice</b>	<b>PriceGrabber-Novice</b>	<b>Yahoo-Novice</b>
		176.4	163.5	241.2
<b>Clusty-Novice</b>	176.4		P = 0.6061	P = 0.0328
<b>PriceGrabber- Novice</b>	163.5	P = 0.6061		P = 0.0120
<b>Yahoo-Novice</b>	241.2	P = 0.0328	P = 0.0120	

This result shows that in the novice search condition it took less time on PriceGrabber and Clusty Shopping than on Yahoo Shopping, but there was no significant difference between the search duration times on PriceGrabber and Clusty Shopping. Thus, hypothesis 1a was not supported by the result, because participants' search performance on rich-faceted PriceGrabber was significantly better than on clustering-supported Clusty Shopping only for expert searches, but not for novice searches. The search performances of the less-faceted Yahoo Shopping were not significantly better than clustering-supported Clusty Shopping for expert searches, and even lower than the search performance of Clusty Shopping for novice searches.

H1a. Consumers with little product knowledge will perform the search task more efficiently on the shopping search websites with faceted search support than on the shopping search websites with clustering support. (Websites with faceted search support: system A and B)

H1b. Consumers with little product knowledge will perform the search task more efficiently on the shopping search website with rich faceted search support than on the shopping search website with less faceted search support.

The hypothesis 1b was supported by the results, because participants' search performance on rich-faceted PriceGrabber was significantly better than on less-faceted Yahoo Shopping for novice searches. PriceGrabber performed better than Yahoo Shopping for expert searches as well.

### **Perceived Values: Usefulness and Directability/Controllability**

#### **Usefulness**

Participants' perceived usefulness assessments for each website were different among the three websites. In expert searches, there was a statistically significant difference among the three websites: Perceived usefulness of PriceGrabber was statistically significantly higher than that of Yahoo Shopping, which in turn was statistically significantly higher than that of Clusty Shopping.

Table3. Perceived usefulness for "expert" searches (on 7 point Likert scale; Mean)

<b>Clusty-Expert</b>	<b>PriceGrabber-Expert</b>	<b>Yahoo-Expert</b>
3.14	5.78	4.67

\*ANOVA test:  $P < 0.001$ , Sig. level = 0.05

		<b>Clusty-Expert</b>	<b>PriceGrabber-Expert</b>	<b>Yahoo-Expert</b>
		3.14	5.78	4.67
<b>Clusty-Expert</b>	3.14		$P < 0.0001$	$P < 0.0001$
<b>PriceGrabber-Expert</b>	5.78	$P < 0.0001$		$P = 0.0004$
<b>Yahoo-Expert</b>	4.67	$P < 0.0001$	$P = 0.0004$	

In novice searches, there also was a statistically significant difference among the three websites.

Table4. Perceived usefulness for “novice” searches (on 7 point Likert scale; Mean)

<b>Clusty-Novice</b>	<b>PriceGrabber-Novice</b>	<b>Yahoo-Novice</b>
3.53	5.94	3.92

\*ANOVA test:  $P < 0.001$ , Sig. level = 0.05

		<b>Clusty-Novice</b>	<b>PriceGrabber-Novice</b>	<b>Yahoo-Novice</b>
		3.53	5.94	3.92
<b>Clusty-Novice</b>	3.53		$P < 0.0001$	$P = 0.2873$
<b>PriceGrabber-Novice</b>	5.94	$P < 0.0001$		$P < 0.0001$
<b>Yahoo-Novice</b>	3.92	$P = 0.2873$	$P < 0.0001$	

Perceived usefulness of PriceGrabber was statistically significantly higher than Yahoo Shopping and Clusty Shopping, but there was no statistically significant difference between Yahoo Shopping and Clusty Shopping.

### **Directability/Controllability**

How well the participants thought that they were able to direct their searches on each website for each product was measured by 7 point Likert scales through the questionnaire. According to their answers, their perceived directability/controllability for each website was different among the three websites. In expert searches, there was a statistically significant difference regarding the directability among the three websites. The perceived directability of PriceGrabber was statistically significantly higher than

Yahoo Shopping, and the one on Yahoo Shopping was statistically significantly higher than Clusty Shopping.

Table5. Perceived directability for “expert” searches (on 7 point Likert scale; Mean)

<b>Clusty-Expert</b>	<b>PriceGrabber-Expert</b>	<b>Yahoo-Expert</b>
3.64	5.81	4.83

\*ANOVA test:  $P < 0.001$ , Sig. level = 0.05

		<b>Clusty-Expert</b>	<b>PriceGrabber-Expert</b>	<b>Yahoo-Expert</b>
		3.64	5.81	4.83
<b>Clusty-Expert</b>	3.64		$P < 0.0001$	$P = 0.0009$
<b>PriceGrabber-Expert</b>	5.81	$P < 0.0001$		$P = 0.0019$
<b>Yahoo-Expert</b>	4.83	$P = 0.0009$	$P = 0.0019$	

In novice searches, there was also a statistically significant difference regarding directability among the three websites.

Table6. Perceived directability for “novice” searches (on 7 point Likert scale; Mean)

<b>Clusty-Novice</b>	<b>PriceGrabber-Novice</b>	<b>Yahoo-Novice</b>
3.56	5.97	4.00

\*ANOVA test:  $P < 0.001$ , Sig. level = 0.05

		<b>Clusty-Novice</b>	<b>PriceGrabber-Novice</b>	<b>Yahoo-Novice</b>
		3.56	5.97	4.00
<b>Clusty-Novice</b>	3.56		$P < 0.0001$	$P = 0.2288$
<b>PriceGrabber-Novice</b>	5.97	$P < 0.0001$		$P < 0.0001$
<b>Yahoo-Novice</b>	4.00	$P = 0.2288$	$P < 0.0001$	

Perceived directability of PriceGrabber was statistically significantly higher than Yahoo Shopping and Clusty Shopping, but there was no statistically significant difference between Yahoo Shopping and Clusty Shopping.

These results partly supported hypothesis 2a. The perceived values (usefulness and directability) of the rich-faceted search system, PriceGrabber, were significantly higher than the clustering-supported system, Clusty Shopping, for both expert and novice searches. However, the perceived value of the less-faceted system, Yahoo Shopping, was significantly higher than the value of Clusty Shopping only for expert searches.

H2a. Consumers with little product knowledge will find the search experience at the faceted search websites more useful and controllable than at the search website with clustering support.

The hypothesis 2b was fully supported by these results since the perceived values of rich-faceted search system, PriceGrabber, were significantly higher than the values of less-faceted system, Yahoo Shopping for both expert and novice searches.

H2b. Consumers with little product knowledge will find the search experience at the rich-faceted search websites more useful and controllable than at the less-faceted search website.

### Within Websites

Next, each search website's perceived usefulness and directability in expert searches and novice searches were compared.

Table7. Perceived usefulness of each website for expert/novice searches (on 7 point Likert scale; Mean)

	<b>Clusty Shopping</b>	<b>PriceGrabber</b>	<b>Yahoo Shopping</b>
<b>Expert search</b>	3.14	5.78	4.67
<b>Novice search</b>	3.53	5.94	3.92
<b>Difference between Expert vs. Novice search (Sig. level=0.05)</b>	P = 0.2895	P = 0.5523	P = 0.0347

There were no significant differences found between expert searches and novice searches on PriceGrabber and Clusty Shopping, but on Yahoo Shopping the difference was found to be statistically significant. Perceived usefulness of Yahoo Shopping significantly decreased when participants conducted novice searches ( $P=0.0347$ , Sig.=0.05). On the other hand, although not statistically significant, even the mean scores of perceived usefulnesses of both PriceGrabber and Clusty Shopping were higher with novice searches than with expert searches.

In terms of perceived directability, there were no statistically significant differences between expert searches and novice searches on PriceGrabber or Clusty Shopping. But on Yahoo Shopping, participants' perceived directability with novice

searches was statistically significantly lower than the one with expert searches, as shown in Table8 ( $P=0.0195$ , Sig.=0.05).

Table8. Perceived directability of each website for expert/novice searches (on 7 point Likert scale; Mean)

	<b>Clusty Shopping</b>	<b>PriceGrabber</b>	<b>Yahoo Shopping</b>
<b>Expert search</b>	3.64	5.81	4.83
<b>Novice search</b>	3.56	5.97	4.00
<b>Difference between Expert vs. Novice search (Sig. level=0.05)</b>	$P = 0.8186$	$P = 0.5317$	$P = 0.0195$

#### **Likelihood of Future Use**

Also, participants answered the questions about their likelihood of using each website for their future online shopping searches. PriceGrabber had the highest score, followed by Yahoo Shopping and then Clusty Shopping. Significant differences were found among the scores of all three search systems.

Table9. Likelihood of using the website for future shopping (on 7 point Likert scale; Mean)

<b>Clusty Shopping</b>	<b>PriceGrabber</b>	<b>Yahoo Shopping</b>
2.36	5.61	4.08

\*ANOVA test:  $P < 0.001$ , Sig. level = 0.05

## **Further Comparison of Performance and Perceived Values**

### **Performance and Likelihood of Future Use**

Correlation between participants' search duration time on each website for expert/novice searches and their answers regarding the likelihood of future use was tested. For most of the pairs, there were no statistically significant correlations. However, search duration time for novice searches on PriceGrabber showed a moderate negative correlation with the likelihood of future use of the website ( $-.438$ ,  $\text{Sig.}=0.008$ ).

Since no relevant correlations were found between search performances and participants' likelihood of future use of the website, with the exception of PriceGrabber for novice searches, hypothesis 3 was not supported. However, for novice searches on PriceGrabber, there was a very relevant ( $\text{Sig.} = 0.008$ ) correlation between the search performance and the likelihood of future use in a moderate power ( $-.438$ ).

H3. Search performances for finding novice products will have a greater impact on consumers' willingness to use the website in the future than search performances for finding expert products.

### **Perceived Usefulness and Likelihood of Future Use**

Correlation between participants' perceived usefulness of each website for expert/novice searches and their answers to the likelihood of future use question was also tested. The results showed positive correlations between the perceived usefulness of all three websites with both expert and novice searches and the likelihood of future use of the websites.



Table10. Correlation between perceived usefulness and likelihood of future use

	<b>Expert</b>	<b>Novice</b>
<b>Usefulness_C</b>	.416*	.623**
<b>Usefulness_P</b>	.424**	.662**
<b>Usefulness_Y</b>	.637**	.672**

### **Perceived Directability and Likelihood of Future Use**

The relationships between the perceived directability and the likelihood of future use also showed positive correlations for all three websites, with the exception being that the correlation between the directability of PriceGrabber in expert searches and the likelihood of future use of PriceGrabber was low and was not statistically significant ( $r = .160$ , Sig. = .351).

Table11. Correlation between perceived directability and likelihood of future use

	<b>Expert</b>	<b>Novice</b>
<b>Directability_C</b>	.367*	.545**
<b>Directability_P</b>	.160	.621**
<b>Directability_Y</b>	.475**	.660**

Overall, both the perceived usefulness and the perceived directability showed positive correlations with the likelihood of future use of the websites. The correlations were consistently stronger in novice searches than in expert searches. Thus, it can be claimed that hypothesis 4 was supported by the results, since stronger correlations were consistently found in the relationships between the perceived values and the likelihood of future use of the websites in novice searches rather than in expert searches.

H4. Consumers' perceived values (usefulness and directability/controllability) regarding the websites in novice product searches will have a greater impact on consumers' willingness to use the websites in the future than their perceived values of the websites in expert product searches.

## **Discussion**

The purpose of this study was to investigate two research questions on online shoppers' product search behavior on three shopping search websites using different types of search tools: Firstly, search performance (search duration time) and secondly, user experience (usefulness, directability) when online shoppers use a rich-faceted search (PriceGrabber), a less-faceted search (Yahoo Shopping), and a clustering search (Clusty Shopping).

Regarding the search performance of the three websites in the expert search condition, PriceGrabber was superior to the other two websites, and there were no statistically significant differences between the performances of Yahoo Shopping and Clusty Shopping. But in novice searches, both PriceGrabber and Clusty Shopping yielded statistically significantly shorter search duration times than Yahoo Shopping, while there was no statistically significant difference between PriceGrabber and Clusty Shopping. These results may imply that rich-faceted searches perform better than less-faceted searches in both expert searches and novice searches. Users would spend significantly less time finding the products that they are looking for in rich-faceted search systems than less-faceted search systems, no matter whether they search for items that they know well or that they do not know much about. It implies that just using faceted search tools is not

enough to help consumers find products, but providing various facets for users to find relevant components for their search result-refining purposes would be more beneficial to the online consumers. When designing facets for certain product items, consumers' specific needs in product searching and search refinement would also have to be considered.

The results may also imply that the cost-effective, clustering search's performance is not significantly different from the performance of the less-faceted search in expert search conditions. These results also suggest that a clustering search's performance may even be better than a less-faceted search and as effective as a rich-faceted search in novice search conditions. Conversely, this would indicate that a less-faceted search is not as effective as a clustering search in novice search conditions. In addition, the fact that the less-faceted search did not outperform the clustering search in both expert and novice search conditions could indicate that the amount of effort invested on constructing a faceted search system may not yield commensurate levels of returns if the facets are not rich, many, or relevant enough to support users' search refinements.

For both expert searches and novice searches, the rich-faceted PriceGrabber was rated the highest in perceived usefulness. The less-faceted Yahoo Shopping placed second in expert searches, but not in novice searches. In novice searches, the score of Yahoo Shopping dropped and did not show statistically significant differences when compared to the clustering-supported Clusty Shopping. It implies a decrease in usefulness of Yahoo Shopping in novice search conditions. The usefulness scores in expert search conditions and in novice search conditions were not statistically significantly different on PriceGrabber and Clusty Shopping. However, on Yahoo Shopping the scores

significantly decreased in the novice search condition. Lack of relevant facets for the search tasks on Yahoo Shopping seems to have influenced the users' difficulties in refining the search results and finding the right products in novice searches.

Moving our attention to the directability/controllability of the websites, the rating results were similar to the ones for perceived usefulness. PriceGrabber was superior to the other two search websites, while Yahoo Shopping excels Clusty Shopping in expert searches, but not in novice searches. It may be because in novice searches, participants with little product knowledge did not get enough help for their search refinement on Yahoo Shopping, which employed limited numbers of facets. Also, it is possible that the participants in the novice condition could not form better queries to find the products, which the participants in the expert condition might have formed.

The comparisons of the perceived values—usefulness and directability—in expert and novice search conditions indicate that only Yahoo Shopping had statistically significantly lower scores in the novice search condition than in the expert search condition. However, neither the usefulness nor the directability scores showed significant differences between the expert and novice conditions for PriceGrabber and Clusty Shopping. This result may imply that users' experiences in rich-faceted searches and clustering searches do not significantly worsen in novice searches, but maintain the quality of the experience even when conducting novice searches. However, when conducting novice searches on Yahoo Shopping, where, in these instances, fewer relevant facets are provided, users may not find it useful or have a sense of control and the quality of user experiences may decrease accordingly.

From these results, it could be claimed that a rich-faceted search system may provide top-level performance in terms of search efficiency and quality of the user's experience, compared to a less-faceted search system. Although the perceived values—usefulness and directability—are lower than in the rich-faceted search system, the clustering-supported search system could perform almost as efficiently as the rich-faceted system in novice searches, but not in expert searches. The less-faceted search seems to provide a comparatively better user experience regarding usefulness and directability in expert searches, but not in novice searches. The finding that the less-faceted search system did not statistically significantly excel any of the rich-faceted or clustering-supported search system in search efficiency is noticeable. The less-faceted search system might need to be improved regarding its facets. The facets can be relevant to users' search refining needs and various enough to assist both expert product users and novice product users.

The statistically significant differences concerning the likelihood of future use found between the three websites may imply that users prefer a rich-faceted search system over a less-faceted search system and a faceted search system (no matter whether it is rich- or less-faceted) over a clustering search system.

PriceGrabber (5.61) > Yahoo Shopping (4.08) > Clusty Shopping (2.36) (Mean scores of likelihood of future use)

Participants answered an open question about what features they liked on certain websites. In their answers, most of them indicated that they liked PriceGrabber because its rich facets helped them refine the search results effectively, whereas they found Yahoo

Shopping less useful because the facets did not include some relevant factors to help them refine the search results. Many of them also mentioned that the clusters on the Clusty Shopping website were not very useful, not always efficient, and were even confusing.

The relationship between the search performance and the likelihood of future use was found to have almost no correlation for most of the websites in expert/novice conditions. However, for PriceGrabber's novice searches, a relevant correlation between the search duration time and the likelihood score was found. This could mean that the less time users spend in novice searches, the higher the likelihood of their using PriceGrabber in the future. It could also mean that the impact of search efficiency on users' likelihood of future use of the website was greater in novice searches. PriceGrabber's power, as reflected in its high scores in likelihood of future use, seems to lie significantly in its effectiveness at results-refining, especially in novice searches.

When it comes to the relationship between the usefulness and the likelihood of future use for all three websites for both expert and novice search conditions, the correlations were found to be positive and significant. The correlations between the directability and the likelihood of future use were also found positive and significant in most cases (five out of six cases on three websites in expert/novice conditions). The only exception was that the positive correlation between the directability of PriceGrabber in expert searches and the likelihood of future use was low and not statistically significant. This may imply that consumers would like to use product search websites whose search-refining functions are useful and where they feel they have more of a sense of control or directability over searching and refining.

The correlations between perceived usefulness and likelihood of future use were higher in novice searches than in expert searches for all three websites. There was also a tendency for higher correlations between perceived directability and likelihood of future use in novice search conditions than in expert search conditions.

In summing up the comparisons of the correlations in expert and novice searches, the results indicate that users would like to use product search websites where they can get the most help for searching and refining and where they can direct their searches better, especially when they search for products that they do not have much knowledge or experience with.

From the findings in Jaillet (2003), it can be derived that consumers with little product knowledge are more likely to get help from shopping search websites than expert consumers are. This has implications relevant to the present study, which found that the less-faceted product search website showed significantly lower performance and lower perceived values (usefulness and controllability) than the rich-faceted system in the novice search condition. Further, compared to the clustering-supported system, the less-faceted search system's performance was significantly lower and its perceived values were just comparable (no significant difference found in terms of usefulness and controllability) in novice search conditions. Thus, it is important to improve search-aid tools for novice product searches on shopping websites. In particular, product search websites with faceted search support may want to consider improving and enriching their facets so that novice product consumers may better make use of them. Also, it would be desirable for a product website to have a search interface design that better assists consumers with novice product knowledge.

The rich-faceted search system, PriceGrabber, was found to show significantly higher search performance and perceived values than the clustering-supported search system and the less-faceted search system in expert search conditions. Also, in novice search conditions, PriceGrabber showed relatively superior performance and perceived values than the other search websites, with the exception that it just showed a higher, but not statistically significant, search performance than Clusty Shopping in novice searches. Its rich facets, which can be useful for search refinement, can functionally help consumers find the right products more efficiently with better user experiences, regardless of whether consumers have relevant product knowledge.

The efficiency of a clustering-supported search system in novice search conditions was also supported by some of the findings in this study. However, despite the efficient search performance of the clustering-supported system, its low scores in the perceived values and the likelihood of future use indicate that there are still other issues which may prevent clustering search systems from being more widely accepted and used. Its effectiveness in search performance, which was comparable to the performance of PriceGrabber in novice searches, should be noted. Considering the efficient functionality of clustering search in novice search conditions, it could also be proposed that other shopping websites or even general search websites could benefit from employing or incorporating the clustering function to support search refinement in novice search conditions. Decisions concerning the selection of search tools would depend on the type of website and its users, as well as a comparison of the costs and effects of using a clustering search system or a faceted search system.



### **Advantages and Limitations of the Study**

Since the Internet and the web are open systems that run on a public infrastructure and continuously change according to system builders' choices or other reasons, it is difficult, in a strict sense, to design true controlled lab experiments to test the research hypotheses. Also, it would be fairly difficult to conduct experiments with the same websites that were used in the previous experiments because of these websites' dynamically changing characteristics. However, the decision was made to use actual Internet shopping search websites because observation of real-world Internet environments is vital to understanding people's online behavior and measuring the true performance of different search technologies. On the other hand, it may be to the advantage of this study that the real world shopping search websites and their search tools were used to answer the research questions. For practical purposes, the implications of the study results could be applied to practitioners' research or their work, since this study was based on and actually conducted with some of the active, major shopping search websites.

For the experiment of this study, recruiting advertisements were posted on the campuses of local universities. This may imply two limitations of the sampling of the population. First, the subjects are more likely to be young undergraduate and graduate students. They may be better at Internet searching and browsing than people in other age groups. Thus, it might be difficult to generalize the results of the study to the general public. Also, the subjects are more likely to be the residents of North Carolina in the United States, who might have different experiences in shopping online, than people in other states or other countries with different cultures and online environments. Thus, it

might be also difficult to generalize the results across other states or countries.

Conducting a more extended version of this research might be able to yield better results, if it uses a bigger, randomized sample size and also uses search websites that are custom-made to have all the same features except for the types of search tools.

This study focused on consumers' online product search performance when using certain shopping search websites. Thus, it did not necessarily cover topics and factors regarding online shopping other than search efficiency, users' perception of usefulness, and users' search experiences. This study has a variety of possibilities to be expanded across many different disciplines. For example, considering the possible effects of emotions on people's perception about usability (Norman, 2002), future studies may become more interdisciplinary by including such psychological factors as consumers' positive or negative emotions while conducting searches.

### **Conclusions**

As more and more people, including merchants and consumers, go online with the continuously growing supplies and demand in online markets, online shopping search websites are flooded by numerous numbers of products and stores which require more searching efforts and search costs to online consumers. Thus, helping to improve users' product search performance and experience in this fast growing e-commerce environment is an important issue at this time.

The present research examined the search performance of shopping search engines in terms of their effectiveness for consumers with little product knowledge to find the product that they wanted to purchase. Also, this research investigated users'

perceptions of usefulness and directability in their searches as indicators of quality of the product search experience.

PriceGrabber, which was judged to have richer facets for the search tasks of the study, not only performed better in search effectiveness in most cases, but also excelled when compared to the other two types of search websites in terms of users' perceived usefulness, directability/controllability, and likelihood of future use. Clustering-supported Clusty Shopping showed efficient search performance in novice searches, which was comparable to the search performance of PriceGrabber, but was not found to be favorable to users in terms of its perceived usefulness, controllability, and likelihood of future use. Yahoo Shopping, which was regarded as having fewer numbers of relevant facets for the search tasks, showed relatively weaker performances in novice searches than the other two types of search systems. Further, in novice searches, less-faceted Yahoo Shopping was not only perceived to be less useful and less controllable than a rich-faceted search system, but was also perceived to have only a comparable level of usefulness and controllability to the automatic, clustering search system.

The research findings of this study contribute to enhancing the understanding of effective user interface for online shopping websites and suggest implications for the actual design of online search tools for practical use. Potential groups of people who may be interested in these findings include researchers in related interdisciplinary fields such as human-computer interaction and information retrieval (those concerned with search performance and user interfaces) and e-commerce and marketing science (those concerned with consumer information processing and the decision-making process).

Also, the findings of this research are useful to the practitioner groups in the IT/search engine industry.

In sum, this paper presented a comparison of three different search tools in expert and novice search conditions. The themes of this study incorporate interests and perspectives from several different disciplines, such as information and library science, human-computer interaction, psychology, and business administration. Additionally, although not all of the different perspectives and research interests were covered in this study in appropriate depth, this study is another step toward more interdisciplinary research regarding online system usability and consumer behavior. Lastly, it is the author's hope that the topics of this study will be further developed through the collaboration of researchers and practitioners from different fields, contributing to the development of related research areas and actual search systems, and ultimately enriching human behaviors online.

## References

- Bruza, P. D., Dennis, S., & McArthur, R. (2000). Interactive internet search: Keyword directory and query reformulation mechanisms compared. *Proceedings of SIGIR'2000*, 280-287.
- DeSarbo, W. S., & Choi, J. (1999). A latent structure double hurdle regression model for exploring heterogeneity in consumer search patterns. *Journal of Econometrics*, 89, 423-455.
- Haubl, G., & Trifts, V. (2000). Consumer decision making in online shopping environments: The effects of interactive decision aids. *Marketing Science*, 19(1), 4–21.
- Hearst, M. (2006, August). Design recommendations for hierarchical faceted search interfaces. *Proc. ACM SIGIR Workshop on Faceted Search, Seattle, WA* (pp. 26-30).
- Hearst, M. A. (2006). Clustering versus faceted categories for information exploration. *Communications of the ACM*, 46(4), 59–61.
- Hsieh-Yee, I. (1993). Effects of search experience and subject knowledge on the search tactics of novice and experienced searchers. *Journal of the American Society for Information Science*, 44(3), 161-174.
- Jaillet, H. (2003). Web metrics: measuring patterns in online shopping. *Journal of Consumer Behaviour*. 2(4), 369-381.

- Jansen, B. J., Spink, A., & Taksa, I. (2009). *Handbook of research on web log analysis*. Hershey, PA: IGI Global.
- Kumar, N., Lang, K.R., & Peng, Q. (2005). Consumer Search Behavior in Online Shopping Environments. In Proceedings of the 38th Hawaii International Conference on System Sciences (HICSS-38), January.
- Kumar, N., & Lang, K.R. (2007). Do search terms matter for online consumers? *Decision Support Systems*. 44(1), 159-174.
- Norman, D. A. (2002). Emotion and design: Attractive things work better. *Interactions Magazine*, ix (4), 36-42.
- Smith, M., Bailey, J., & Brynjofsson, E. (1999). Understanding digital markets: Review and assessment, In Erik Brynjolfsson and Brian Kahin,(Eds), *Understanding the Digital Economy: data, tools, and research* (99-136). Cambridge, MA:MIT Press.
- Tullis, & Albert. (2008). *Measuring the User Experience* (19). Burlington, MA: Morgan Kaufmann.
- Wildemuth, B.M., (2004). The effects of domain knowledge on search tactic formulation. *Journal of the American Society for Information Science and Technology*, 55, 246–258.
- Wildemuth, B.M. (2009). *Applications of Social Research Methods to Questions in Information and Library Science*. Wesport, CT: Libraries Unlimited.

## Appendices

**Appendix A – Experiment instruction material (used for the computer group in the Yahoo Shopping-PriceGrabber-Clusty Shopping order.)**

### Search Project 1 - Hard drive

Imagine that you want to purchase an external hard drive to regularly backup your data in your notebook PC.

**Find “hard drive” that is portable with these conditions:**

- Capacity: **500GB or larger**
- Type: **2.5” slim size**
- Price: **Less than \$130**
- When you have found at least 5 items described above, click the “Task-Done” bookmark on the web browser.

• **Task order:**

1. Begin your search on  by clicking on the bookmark placed at the top of your browser and follow the instructions below.

- 1) Start the search by entering the keywords “hard drive” for the product.
- 2) Refine search results by using the search-aid feature on the left side of the website.
- 3) When you have found an appropriate item, draw circle on the product item with your mouse cursor.
- 4) If you can’t find enough items by refining the search result, you may enter modified keywords in the search box to find the right products.
- 5) When you have found 5 items or more, click on the “Task-Done” bookmark.

2. Now use  to do the same task 1) - 5).

3. Now use  to do the same task 1) - 5).

\* Below are sample product images.



When you are done with all 3 tasks, please call the investigator before moving to the next project.

## Search Project 2 – Musical Keyboard

Imagine that you haven't played piano for a long time and you want to buy a portable keyboard to get back into it.

Find **"Musical Keyboard"** with these conditions:

- Number of keyboard keys: **62 keys or more (covering 6 octaves or more)**
- Type: **Portable keyboards (not "digital pianos")**
- Price: **Less than \$700**
- When you have found at least 5 items described above, click "Task-Done" button on the browser.

### • Task order:

1. Begin your search on  by clicking on the bookmark placed at the top of your browser.

- 1) Start the search by entering the keywords "musical keyboard" for the product.
- 2) Refine search results by using the search-aid feature on the left side of the website.
- 3) When you have found an appropriate item, draw circle on the product item with your mouse cursor.
- 4) If you can't find enough items by refining the search result, you may enter modified keywords in the search box to find the right products.
- 5) When you have found 5 items or more, click on the "Task-Done" bookmark.

2. Now use  to do the same task 1) - 4).

3. Now use  to do the same task 1) - 4).

\* Below are sample product images.

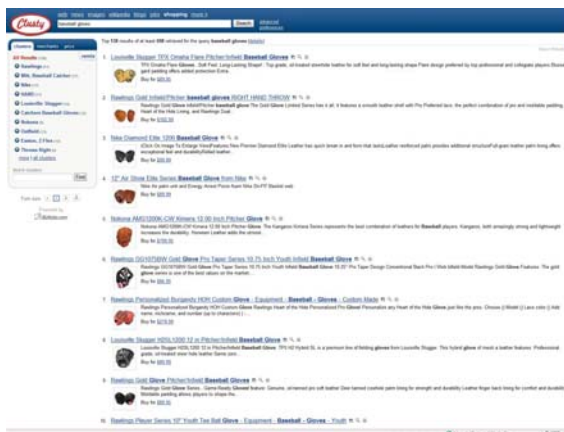


When you are done with all 3 tasks, please call the investigator.

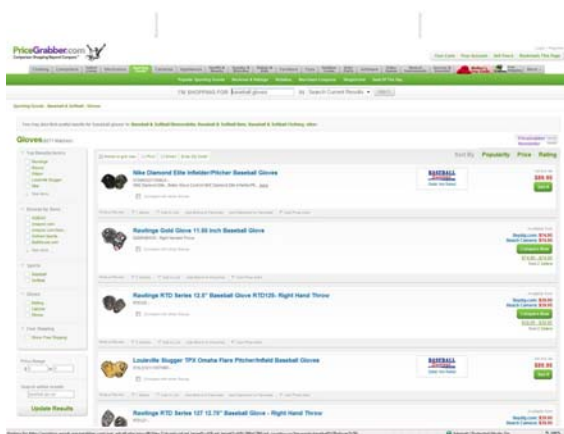


## Appendix B – Product search websites

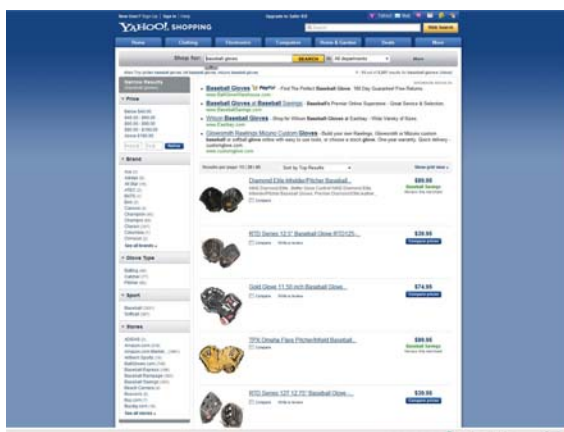
Clusty Shopping (<http://shopping.clusty.com/>)



PriceGrabber (<http://www.pricegrabber.com/>)



Yahoo Shopping (<http://shopping.yahoo.com/>)



### **Acknowledgements**

I would like to express my sincere gratitude to everyone who has supported me throughout the years of my master's study. Thank you all for your help, support, and guidance.

I would like to begin by expressing my deepest appreciation to my advisor, Dr. Gary Marchionini. Thank you for teaching and guiding me from the beginning of the program to the writing of this master's paper, leading me throughout my joyful exploration of HCI and Information Science at SILS. Dear Dr. Evelyn H. Daniel, my mentor and life advisor here at SILS, I can't thank you enough for your warm heart and great support. Had it not been for your encouragement, I would not have applied to and become a CRADLE fellow. I would like to thank Dr. Hemminger, my CRADLE advising professor, for his great support and guidance in learning new skills and working on innovative projects.

The experiments of this study were done through SILS with the support from the Carnegie Grants. I would like to thank the School of Information and Library Science at UNC-Chapel Hill, its staff, professors, and colleagues for the academically excellent and supportive environment.

I would also like to thank my colleagues and friends, Luke Miller, Lina Huang, Jessica Lin, Yaxiao Song, Jung Sun Oh, Sanghee Oh, and all my friends and colleagues at SILS for their great help, support, and friendship. Especially, I would like to thank my friends Steven Bongkyun Chun, Kristin Fleetwood, Julia TerMaat, and Sarah Park, for their kind help and great advice in editing my writing. Also, my deep gratitude goes to my brothers and sisters at the Duraleigh Presbyterian Church in Raleigh, North Carolina and the Abundant Life Church in Bundang, South Korea, for their love, support, and constant prayer.

Lastly, special sincere thanks go to my parents, Taeyong Oh and Hyojung Kim, and my brother, Dae Yeol Oh, for their love, prayer, and unfailing support.

Overall, it is through grace that I was able to pursue my studies and research at SILS, UNC-Chapel Hill.